

Official Action. Reconsideration of this application in light of these remarks, and allowance of this application are respectfully requested.

I. Rejection of Claims Under 35 U.S.C. § 103(a)

Claims 1 and 7 were rejected under § 103(a) as unpatentable over U.S. Patent No. 5,463,618 to Furukawa et al. (hereinafter, Furukawa) in view of Applicants' admitted prior art. According to the Examiner, Furukawa discloses "an echo canceller with Normalized Least Mean Square algorithm[:] Determining a pseudo acoustic signal[:] Providing for holding signals[:] Subtraction of the pseudo acoustic echo signal[: and] Sampling the input signal at 8kHz." (December 31, 2002 Official Action at page 2.) The Examiner later provides that Furukawa, "describes the functionality of the echo canceller with implementation of first and second adaptive filters which update the filter coefficients by NLMS and a voice detector which reads on 'control means. . . ' and 'control step. . . ' (Id. at page 8.) Applicants disagree with the Examiner's characterization of the prior art, and for the reasons stated below, traverses the Examiner's rejections.

The present invention, as recited in amended claim 1, is directed to a speech processing apparatus comprising *inter alia*: a decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information of said echo-canceled signal; and a control means for, in a frame for which the result of a decision made by said decision means is negative, storing in said storage means the current impulse response held by said supply means and, in a frame for which the result of the decision is

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

positive, retrieving one of the impulse responses stored in said storage means and supplying it to said supply means. Claim 7 is directed to a speech processing method that recites steps corresponding to the limitations recited in claim 1. According to Applicants' specification:

When the above decision process decides that the voice decision flag is OFF, a control means aa6 retrieves the current impulse response from the supply means aa7 and stores it as the desired impulse response in the storage means aa2. When the voice decision flag is ON, there is a possibility of the impulse response held in the supply means aa7 having deviated from a desired value, so that the control means aa6 retrieves one of the impulse responses stored in the storage means and overwrites the impulse response held in the supply means aa7 with the retrieved one.

(Applicants' Specification at page 53, lines 11-24.) (Emphasis added.) Later the Applicants repeat the process when they provide that "where the VAD [voice activity detection] is ON, the filter coefficient stored at the m-th location of the filter coefficient buffer is retrieved and the degraded filter coefficient is reset by the retrieved value. " (Applicants' Specification at page 33 lines 17-20.) In making the various references to the specification set forth herein, it is to be understood that Applicants are in no way intending to limit the scope of the claims to the exemplary embodiments shown in the drawings and described in the specification. Rather, Applicants expressly affirm that they are entitled to have the claims interpreted broadly, to the maximum extent permitted by statute, regulation and applicable case law.

To establish a prima facie case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be some reasonable expectation of success. Finally, the prior art references must teach or suggest all the

claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, not in Applicants' disclosure.

In this case, the prior art references fail to teach or suggest at least three of the claimed elements. First, none of the cited prior art teaches, suggests or discloses at least "a control means for . . . in a frame for which the result of [the] decision is positive, retrieving one of the impulse responses stored in said storage means and supplying it to said supply means." In the Official Action, the Examiner contends that Furukawa teaches this feature, and to support her position she cites Col. 5, line 64 – Col. 6, line 12 of Furukawa, which provides that:

the first and second adaptive filters 1 and 3 renew their filter coefficients by NLMS (Normalized Least Mean Square). In FIG. 1, the echo canceller further includes a divider 7 for dividing the second ratio R2 by the first ratio R1 (hereinafter, $R2/R1=R3$ is referred to as "a third ratio"), and a voice detector 8 for detecting short time power of a received input signal to detect whether far-end speech is present or absent, where the second adaptive filter 3 executes adaptation when the voice detector 8 has detected a far-end speech. In the echo canceller, a double talk detector 9 controls to estimate an impulse response of an echo path of the first adaptive filter 1 when, as a result of the voice detector 8 having detected far-end speech, either one of the first and second ratios R1 and R2 is greater than a first threshold Th1 or the third ratio R3 is greater than a second threshold Th2 (Th2=2, fixed in the embodiments).

(Furukawa at col. 5, line 64 – col. 6, line 12) (Emphasis added.) A careful reading of this passage reveals that the control means in Furukawa only addresses the situation that arises when "the decision is positive" (i.e., when the voice detector has detected speech). An even closer reading reveals that a positive decision is reached when the voice detector in Furukawa has detected a far-end speech. In contrast, a decision is

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

positive in the present invention when the voice detector has detected voice in the microphone input signal (i.e., near-end speech).

Even if the voice detector in Applicants' device and the voice detector in Furukawa detected speech originating from the same vantage point, Furukawa provides that " the second adaptive filter 3 executes adaptation when the voice detector 8 has detected a far-end speech." Adaptation, as discussed in Furukawa, is described as follows:

A comparator 61 compares the first threshold Th1 with the first ratio R1 and yields an output of '1' when $R1 \geq Th1$ and otherwise an output of '0'. A comparator 62 compares the first threshold Th1 with the second ratio R2 and yields an output of '1' when $R2 \geq Th1$ and otherwise an output of '0'. A comparator 63 compares the third ratio R3 with the second threshold Th2 and yields an output of '1' when $R3 \geq Th2$ and otherwise an output of '0'. An OR circuit 64 determines a logical OR among outputs of the comparators 61, 62, and 63. An AND circuit 65 determines a logical AND between an output VD of the voice detector 8 and an output of the OR circuit 64, where adaptation of the first adaptive filter 1 in FIG. 1 is executed when the output ADP of the AND circuit 65 is '1' while the adaptation is suspended when the output is '0'.

(Furukawa at col. 9, lines 1-15.) The present invention in contrast provides that:

When the voice decision flag is ON, there is a possibility of the impulse response held in the supply means aa7 having deviated from a desired value, so that the control means aa6 retrieves one of the impulse responses stored in the storage means and overwrites the impulse response held in the supply means aa7 with the retrieved one.

(Applicants' Specification at page 53, lines 11-24.) Even a cursory reading of the text from Furukawa reveals that the process performed when speech is detected (i.e., adaptation) is not "retrieving one of the impulse responses stored in said storage means and supplying it to said supply means," as recited in Applicants' claims.

In summary, Furukawa does not detect near-end speech, it detects far-end speech, and when the device in Furukawa detects speech, it does not retrieve one of

the impulse responses stored in a storage means and supply it to the supply means. Instead, the device in Furukawa performs a complicated process of adaptation.

Second, the system of Furukawa comprises a first FIR filter and a second filter. The first FIR filter generates a pseudo echo. The second FIR filter determines whether the change of the input signal is a change of echo path or Double Talk. Further, a Double Talk detector controlling the first FIR filter for generating the pseudo echo has a complicated structure using:

- (a) a power calculated by a voice detector on the basis of the signal from the farthest end,
- (b) a ratio of a power of the signal after the first FIR filter executed echo canceling and a power of the signal from the nearest end and
- (c) a ratio of the above two ratios.

On the other hand, the present invention only has the first FIR filter for generating a pseudo echo as shown in Fig. 9. The present invention does not have the second FIR filter. Further, the Double Talk detector uses only a power of a signal after echo canceling. Therefore, the structure of the present invention is simpler than Furukawa and the present invention is different from Furukawa. The present invention realizes a high efficient echo canceller in situations in which additive noises are present. Furukawa does not work in circumstances in which additive noises are present.

Third, in the present invention, the action of Double Talk detection, controlling a FIR filter by the Double Talk detection, inputting/output to/from a buffer of filter coefficient are executed with a frame unit and other processing is performed with a

sample unit as shown in Fig. 9 and at lines 10-15 on page 34. Furukawa, on the other hand, executes all processing at every sample.

Therefore, as previously stated, Furukawa fails to teach, disclose or suggest all of the claim limitations.

Claims 2 and 8 were rejected under § 103(a) as unpatentable over Furukawa in view of applicant's admitted prior art, as applied to claims 1 and 7 above, and further in view of U.S. Patent No. 5,475,791 to Schalk (hereinafter, Schalk). According to the Examiner, neither Furukawa nor Applicants' admitted prior art specifically teach that the echo-cancelled signal is used for speech recognition, and she cites Schalk in an attempt to supply a teaching of this feature. Schalk fails to make up for the shortcomings of Furukawa in view of Applicants' admitted prior art. In other words, the Examiner does not argue and Schalk does not teach a control means as recited in claims 1 and 7. Since claims 2 and 8 depend from claims 1 and 7, respectively, the rejection of claims 2 and 8 under § 103(a) as unpatentable over Furukawa in view of Applicants' admitted prior art, and further in view of Schalk is improper.

The Examiner next rejected claims 3 and 9 under 35 U.S.C. § 103(a) as unpatentable over Furukawa, Applicants' admitted prior art, and Schalk as applied to claims 2 and 8 above, and further in view of "Continuous Speech Recognition in Noise Using Spectral Subtraction and HMM Adaptation," to Flores et al. (hereinafter, Flores). The Examiner admits that neither Furukawa, Applicants' admitted prior art, nor Schalk specifically teach determining a spectrum mean and subtracting the spectrum mean from the spectrum, and she cites Flores for allegedly teaching this feature. Flores does not disclose a control means as recited in claims 1 and 7. Therefore, it does not make

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

up for the shortcomings of the Furukawa, Applicants' admitted prior art, and Schalk combination. Claims 3 and 9 depend indirectly from claims 1 and 7, respectively. Therefore, claims 3 and 9 are not unpatentable under § 103(a) in view of Furukawa, Applicants' admitted prior art, and Schalk as applied to claims 2 and 8 above, and further in view of Flores.

The Examiner next rejected claims 4-5 and 10-11 under 35 U.S.C. § 103(a) as unpatentable over Furukawa, Applicants' admitted prior art, and Schalk as applied to claims 2 and 8 above, and further in view of "Signal Conditioning Techniques for Robust Speech Recognition," to Rahim et al. (hereinafter, Rahim). The Examiner cites Rahim for teaching *inter alia*, that cepstral mean subtraction (CMS) is widely used in speech recognition, and that it improves the robustness in speech recognition by minimizing distortion on the input signal to the recognizer. (December 31, 2002 Official Action at page 6.) Rahim does not make up for the shortcomings of the previous references. Therefore, claims 4-5 and 10-11 are not unpatentable under § 103(a) in view of Furukawa, Applicants' admitted prior art, and Schalk as applied to claims 2 and 8 above, and further in view of Rahim, Flores, and other prior art.

Even though the cited references fail to reach the teachings of Applicants' device, Applicants have amended claims 1 and 7 to more appropriately describe Applicants' invention. Applicants contend that the claims as amended, still patentably distinguish over the prior art. Therefore, the rejection of independent claims 1 and 7 under 35 U.S.C. §103(a) as obvious over some combination of Furukawa, Applicants' admitted prior art, Schalk, Rahim, Flores, and other prior art should be withdrawn. The rejection of dependent claims 2-5 and 8-11 should also be withdrawn as they depend on

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

allowable subject matter as recited in the respective independent claims from which they directly or indirectly depend.

In paragraph 5 of the Official Action, the Examiner rejected claims 6 and 12 under § 103(a) as unpatentable over Rahim and other well-known prior art. According to the Examiner, "Rahim et al. teach determining a cepstrum, calculating the average cepstrum and subtracting the average cepstrum from the cepstrum." The Examiner admits that Rahim does not specifically teach that the cepstrum is obtained by performing a Fourier transform on the spectrum, and she admits that it also does not teach implementing cepstral mean subtraction on a non-speech cepstrum.

(December 31, 2002 Official Action at page 7.) Despite these shortcomings, the Examiner states that:

- (1) it is well-known in the art of speech signal processing to perform a Fourier transform on the logarithm of a spectrum to obtain a cepstrum;
- and (2) it is well-known in the art to provide for estimates of non-speech (or noises) in the implementation of a subtraction scheme for noise suppression.

(December 31, 2002 Official Action at page 7.) Furthermore, the Examiner asserts that it would have been obvious to:

- (1) modify the system of Rahim et al. to perform a Fourier transform on a spectrum in order to obtain a cepstrum, as is well known in the art, for the purpose of efficiently canceling multiplicative distortions; and (2) use a CMS algorithm on a speech cepstrum and a non-speech cepstrum to provide an accurate estimate of other sounds or noise, so as to provide more efficient signal enhancement of the input signal to the speech recognizer.

(Id.) In both instances, the Examiner has failed to point out some teaching, suggestion, or motivation found in either Rahim or in the knowledge generally available to one of ordinary skill in the art to modify Rahim to "perform a Fourier transform on a spectrum in

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

order to obtain a cepstrum," as recited in claim 6. Still further, there is no teaching found in Rahim or the knowledge generally available to one of ordinary skill in the art to "use a CMS algorithm on a speech cepstrum and a non-speech cepstrum to provide an accurate estimate of other sounds or noise, so as to provide more efficient signal enhancement of the input signal to the speech recognizer," as recited in claim 6.

It seems as if the Examiner is relying on personal knowledge to conclude that the present invention is obvious. Consequently, the Applicants must traverse the Examiner's conclusions absent a prior art reference to substantiate her findings. The Patent Statutes clearly provide that when a rejection in an application is based on facts within the personal knowledge of an Examiner, the data should be stated as specifically as possible, and the facts must be supported, when called for by the applicant, by an affidavit from the Examiner. Furthermore, the Examiner may only take official notice of facts outside of the record which are capable of instant and unquestionable demonstration as being "well-known" in the art and, if the Applicants traverse such an assertion, the Examiner should cite a reference in support of his or her position. See, MPEP 2144.03.

Because neither Rahim nor the knowledge generally available to one of ordinary skill in the art teaches: (1) modifying the reference to produce the claimed invention; or (2) a reasonable expectation of success in modifying the reference to produce the claimed invention, the Applicants traverse the Examiners findings because she has failed to meet the initial burden of establishing a *prima facie* case of obviousness.

Pursuant to MPEP 2144.03, if the Examiner continues to maintain the § 103(a) rejection of claims 6 and 12 based on the current grounds, Applicants respectfully request that

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

the Examiner provide support for her assertions. More specifically, Applicants respectfully request the Examiner cite a reference in support of her position that it would have been obvious to: (1) modify the system of Rahim to perform a Fourier transform on a spectrum in order to obtain a cepstrum, as is well known in the art, for the purpose of efficiently canceling multiplicative distortions; and (2) use a CMS algorithm on a speech cepstrum and a non-speech cepstrum to provide an accurate estimate of other sounds or noise, so as to provide more efficient signal enhancement of the input signal to the speech recognizer.

II. Conclusion

In view of the foregoing, it is submitted that the cited prior art considered separately or in combination fails to teach or suggest the Applicants' invention. Therefore, it is respectfully asserted that the present application is in condition for allowance and a notice to that effect is respectfully requested. However, if the Examiner deems that any issue remains after considering this response, she is invited to call the undersigned to expedite the prosecution and work out any such issue by telephone.


FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

If any extension of time under 37 C.F.R. § 1.136 is required to obtain entry of this response, and not requested by attachment, such extension is hereby requested. If there are any fees due under 37 C.F.R. § 1.16 or 1.17 that are not enclosed, including any fees required for an extension of time under 37 C.F.R. § 1.136, please charge those fees to our deposit account 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.

By: 
Leonard Smith, Jr.
Reg. No. 45,118

Dated: June 2, 2003

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

APPENDIX TO AMENDMENT OF MARCH 31, 2003

IN THE CLAIMS:

Please amend claims 1 and 7, as follows:

1. (Twice Amended) A speech processing apparatus comprising:

generation means for generating a pseudo acoustic echo signal for each sample based on a current impulse response simulating an acoustic echo transfer path and on a source signal;

supply means for holding the current impulse response for each sample and supplying the current impulse response to said generation means;

elimination means for subtracting said pseudo acoustic echo signal from a [microphone input] near-end speech signal to remove an acoustic echo component and thereby generate an acoustic echo-canceled signal for each sample;

update means for continually updating the impulse response for each sample by using said source signal, said acoustic echo-canceled signal and the current impulse response held by said supply means and for supplying the updated impulse response to said supply means;

decision means for checking, in each frame, whether or not a voice is included in the [microphone input] near-end speech signal, by using time domain information and frequency domain information of said acoustic echo-canceled signal [, wherein the microphone input signal comprises background noise];

storage means for storing one or more impulse responses in each frame; and

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

control means for, in a frame for which the result of a decision made by said decision means is negative, storing in said storage means the current impulse response held by said supply means and, in a frame for which the result of the decision is positive, retrieving one of the impulse responses stored in said storage means and supplying it to said supply means.

7. (Amended) A speech processing method comprising:

a generation step for generating a pseudo acoustic echo signal for each sample based on a current impulse response simulating an acoustic echo transfer path and on a source signal;

a supply step for holding the current impulse response for each sample and supplying the current impulse response to said generation step;

an elimination step for subtracting said pseudo acoustic echo signal from a [microphone input] near-end speech signal to remove an acoustic echo component and thereby generate an acoustic echo-canceled signal for each sample;

an update step for continually updating the impulse response for each sample by using said source signal, said acoustic echo-canceled signal and the current impulse response held by the supply step and for supplying the updated impulse response to said supply step;

a decision step for checking, in each frame, whether or not a voice is included in the [microphone input] near-end speech signal, by using time domain information and frequency domain information of said acoustic echo-canceled signal;

a storage step for storing one or more impulse responses in each frame; and

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com

a control step for, in a frame for which the result of decision made by said decision step is negative, storing in said storage step the current impulse response held by said supply [means] step and, in a frame for which the result of decision is positive, retrieving one of the impulse responses stored in said storage [means] step and supplying it to said supply [means] step.

FINNEGAN
HENDERSON
FARABOW
GARRETT &
DUNNER LLP

1300 I Street, NW
Washington, DC 20005
202.408.4000
Fax 202.408.4400
www.finnegan.com